

2741

862.2720

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

MASAYUKI YAMADA

Appln. No.: 09/263,262

Filed: March 5, 1999

For: SPEECH SYNTHESIS
APPARATUS, CONTROL
METHOD THEREFOR,
AND COMPUTER-READABLE
MEMORY



Examiner: Unassigned

Group Art Unit: Unknown

August 17, 1999

Box Missing Parts
Assistant Commissioner for Patents
Washington, D.C. 20231

CLAIM TO PRIORITY

Sir:

Applicant hereby claims priority under the International Convention and all rights to which he is entitled under 35 U.S.C. § 119 based upon the following Japanese

Priority Application:

10-057249

Japan March 9, 1998

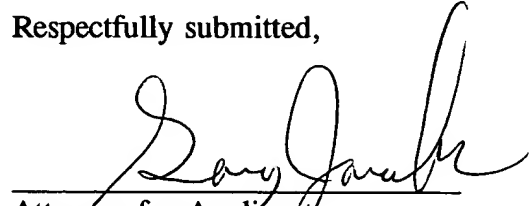
A certified copy of the priority document is enclosed.

K. Ward
8/21/99
#3
Prints
Paper

RECEIVED
AUG 20 1999
TC 2700 MAIL ROOM

Applicant's undersigned attorney may be reached in our Washington, D.C. office by telephone at (202) 530-1010. All correspondence should continue to be directed to our below listed address.

Respectfully submitted,


Attorney for Applicant

Registration No. 28,861

FITZPATRICK, CELLA, HARPER & SCINTO
30 Rockefeller Plaza
New York, New York 10112-3801
Facsimile: (212) 218-2200

09/2103, 2102
Masayuki Yamada
March 5, 1999

(translation of the front page of the priority document of Japanese Patent Application No. 10-057249)



PATENT OFFICE
JAPANESE GOVERNMENT

This is to certify that the annexed is a true copy of the following application as filed with this Office.

Date of Application: March 9, 1998

Application Number : Patent Application 10-057249

Applicant(s) : Canon Kabushiki Kaisha

RECEIVED
AUG 20 1999
TC 2700 MAIL ROOM.

April 2, 1999

Comissioner,
Patent Office

Takeshi ISAYAMA

Certification Number 11-3020564



09/263,262
Masayuki Yamada
March 5, 1999
日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日
Date of Application:

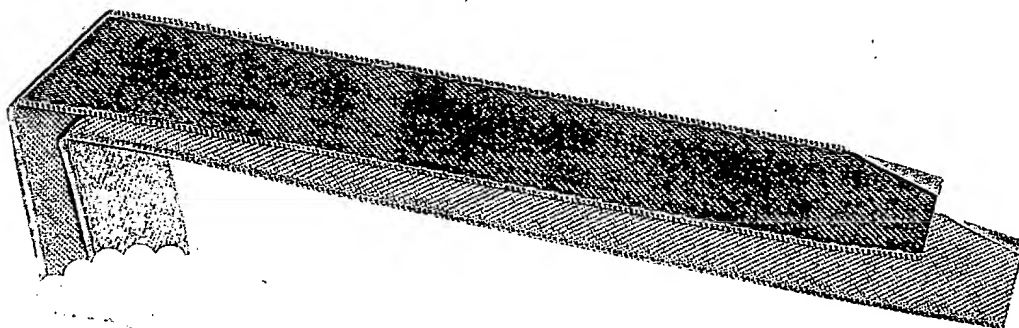
1998年 3月 9日

出願番号
Application Number:

平成10年特許願第057249号

出願人
Applicant(s):

キヤノン株式会社

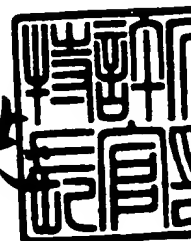


CERTIFIED COPY OF
PRIORITY DOCUMENT

1999年 4月 2日

特許庁長官
Commissioner,
Patent Office

山佐建志



RECEIVED
AUG 20 1999
TC 2700 MAIL ROOM

【書類名】 特許願

【整理番号】 3671016

【提出日】 平成10年 3月 9日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/00

【発明の名称】 音声合成装置及びその制御方法、コンピュータ可読メモリ

【請求項の数】 24

【発明者】

 【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

 【氏名】 山田 雅章

【特許出願人】

 【識別番号】 000001007

 【氏名又は名称】 キヤノン株式会社

【代理人】

 【識別番号】 100076428

 【弁理士】

 【氏名又は名称】 大塚 康德

 【電話番号】 03-5276-3241

【選任した代理人】

 【識別番号】 100093908

 【弁理士】

 【氏名又は名称】 松本 研一

 【電話番号】 03-5276-3241

【選任した代理人】

 【識別番号】 100101306

 【弁理士】

 【氏名又は名称】 丸山 幸雄

【電話番号】 03-5276-3241

【手数料の表示】

【予納台帳番号】 003458

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9704672

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声合成装置及びその制御方法、コンピュータ可読メモリ

【特許請求の範囲】

【請求項 1】 音素片データを管理するデータベースを有する音声合成装置であって、

検索対象の第 1 音素に対し音素環境を考慮した第 2 音素を生成する生成手段と

前記第 2 音素に対応する音素片データを前記データベースより検索する検索手段と、

前記検索手段の検索結果に基づいて、前記音素環境を変更した第 3 音素を生成し、該第 3 音素に対応する音素片データを前記データベースより再度検索する再検索手段と、

前記検索手段あるいは前記再検索手段による検索結果と、前記第 2 音素あるいは前記第 3 音素を対応づけてテーブルに登録する登録手段と

を備えることを特徴とする音声合成装置。

【請求項 2】 前記登録手段は、前記検索手段あるいは前記再検索手段によって検索された音素片データの平均基本周波数を計算する計算手段と、

前記計算手段で計算された平均基本周波数に基づいて、前記検索された音素片データ群を整列する整列手段とを備え、

前記整列手段で整列された音素片データ群の順番で、前記音素片データ群と前記第 2 音素あるいは第 3 音素を対応づけて前記テーブルに登録する

ことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 3】 前記第 2 音素は、前記第 1 音素の左右音素の音素環境を考慮したトライホンである

ことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 4】 前記第 3 音素は、前記第 1 音素の左右音素どちらかあるいはその両方の音素環境を考慮した音素である

ことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 5】 前記第 3 音素は、前記第 1 音素が母音の場合には該第 1 音素

の左音素環境を考慮した音素、子音の場合には該第 1 音素の右音素環境を考慮した音素である

ことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 6】 前記登録手段は、更に、前記検索された音素片データの平均基本周波数を量子化する量子化手段を備える

ことを特徴とする請求項 2 に記載の音声合成装置。

【請求項 7】 前記計算手段は、前記量子化手段で量子化された音素片データ群の各平均基本周波数の内、対応する音素片データが存在しないものについては、その近傍の平均基本周波数で対応する音素片データが存在する平均基本周波数を用いて補間する

ことを特徴とする請求項 6 に記載の音声合成装置。

【請求項 8】 データベースで管理されている音素片データを用いて音声合成を行う音声合成装置であって、

前記データベース中に存在する音素片データの位置を示す位置情報と、該音素片データに対応づけられた音素環境を考慮した音素とを対応づけて管理するテーブルを記憶する記憶手段と、

合成対象の音素群の各音素環境情報とその基本周波数を獲得し、獲得された基本周波数の平均を算出する算出手段と、

前記音素環境情報に対応する音素群を前記テーブルより検索する検索手段と、

前記算出手段で算出された基本周波数の平均に基づいて、前記検索手段で検索された音素群から所定の音素に対応する音素片データの位置情報を前記テーブルより取得する取得手段と、

前記取得手段で取得された位置情報が示す音素片データを前記データベースより取得し、その取得された音素片データの韻律を変更する変更手段と

を備えることを特徴とする音声合成装置。

【請求項 9】 前記変更手段による韻律の変更は、ピッチ同期波形重畳法を用いる

ことを特徴とする請求項 8 に記載の音声合成装置。

【請求項 10】 前記音素環境を考慮した音素の基本周波数が量子化されて

いる場合、前記記憶手段は、その量子化された基本周波数と、該音素に対応する音素片データが存在する前記データベース中の位置を示す位置情報とを対応づけて前記テーブルに管理する

ことを特徴とする請求項 8 に記載の音声合成装置。

【請求項 11】 前記音素環境を考慮した音素の基本周波数が量子化されている場合、前記算出手段は、合成対象の音素群の各音素環境情報を獲得し、また、その量子化された音素群の各基本周波数の平均を算出する

ことを特徴とする請求項 8 に記載の音声合成装置。

【請求項 12】 音素片データを管理するデータベースを有する音声合成装置の制御方法であって、

検索対象の第 1 音素に対し音素環境を考慮した第 2 音素を生成する生成工程と

前記第 2 音素に対応する音素片データを前記データベースより検索する検索工程と、

前記検索工程の検索結果に基づいて、前記音素環境を変更した第 3 音素を生成し、該第 3 音素に対応する音素片データを前記データベースより再度検索する再検索工程と、

前記検索工程あるいは前記再検索工程による検索結果と、前記第 2 音素あるいは前記第 3 音素を対応づけてテーブルに登録する登録工程と

を備えることを特徴とする音声合成装置の制御方法。

【請求項 13】 前記登録工程は、前記検索工程あるいは前記再検索工程によって検索された音素片データの平均基本周波数を計算する計算工程と、

前記計算工程で計算された平均基本周波数に基づいて、前記検索された音素片データ群を整列する整列工程とを備え、

前記整列工程で整列された音素片データ群の順番で、前記音素片データ群と前記第 2 音素あるいは第 3 音素を対応づけて前記テーブルに登録する

ことを特徴とする請求項 12 に記載の音声合成装置の制御方法。

【請求項 14】 前記第 2 音素は、前記第 1 音素の左右音素の音素環境を考慮したトライホンである

ことを特徴とする請求項 12 に記載の音声合成装置の制御方法。

【請求項 15】 前記第 3 音素は、前記第 1 音素の左右音素どちらかあるいはその両方の音素環境を考慮した音素である

ことを特徴とする請求項 12 に記載の音声合成装置の制御方法。

【請求項 16】 前記第 3 音素は、前記第 1 音素が母音の場合には該第 1 音素の左音素環境を考慮した音素、子音の場合には該第 1 音素の右音素環境を考慮した音素である

ことを特徴とする請求項 12 に記載の音声合成装置の制御方法。

【請求項 17】 前記登録工程は、更に、前記検索された音素片データの平均基本周波数を量子化する量子化工程を備える

ことを特徴とする請求項 13 に記載の音声合成装置の制御方法。

【請求項 18】 前記計算工程は、前記量子化工程で量子化された音素片データ群の各平均基本周波数の内、対応する音素片データが存在しないものについては、その近傍の平均基本周波数で対応する音素片データが存在する平均基本周波数を用いて補間する

ことを特徴とする請求項 17 に記載の音声合成装置の制御方法。

【請求項 19】 データベースで管理されている音素片データを用いて音声合成を行う音声合成装置の制御方法であって、

前記データベース中に存在する音素片データの位置を示す位置情報と、該音素片データに対応づけられた音素環境を考慮した音素とを対応づけて管理するテーブルを記憶する記憶工程と、

合成対象の音素群の各音素環境情報とその基本周波数を獲得し、獲得された基本周波数の平均を算出する算出工程と、

前記音素環境情報に対応する音素群を前記テーブルより検索する検索工程と、

前記算出工程で算出された基本周波数の平均に基づいて、前記検索工程で検索された音素群から所定の音素に対応する音素片データの位置情報を前記テーブルより取得する取得工程と、

前記取得工程で取得された位置情報が示す音素片データを前記データベースより取得し、その取得された音素片データの韻律を変更する変更工程と

を備えることを特徴とする音声合成装置の制御方法。

【請求項 20】 前記変更工程による韻律の変更は、ピッチ同期波形重畳法を用いる

ことを特徴とする請求項 19 に記載の音声合成装置の制御方法。

【請求項 21】 前記音素環境を考慮した音素の基本周波数が量子化されている場合、前記記憶工程は、その量子化された基本周波数と、該音素に対応する音素片データが存在する前記データベース中の位置を示す位置情報とを対応づけて前記テーブルに管理する

ことを特徴とする請求項 19 に記載の音声合成装置の制御方法。

【請求項 22】 前記音素環境を考慮した音素の基本周波数が量子化されている場合、前記算出工程は、合成対象の音素群の各音素環境情報を獲得し、また、その量子化された音素群の各基本周波数の平均を算出する

ことを特徴とする請求項 19 に記載の音声合成装置の制御方法。

【請求項 23】 音素片データを管理するデータベースを有する音声合成装置の制御のプログラムコードが格納されたコンピュータ可読メモリであって、

検索対象の第 1 音素に対し音素環境を考慮した第 2 音素を生成する生成工程のプログラムコードと、

前記第 2 音素に対応する音素片データを前記データベースより検索する検索工程のプログラムコードと、

前記検索工程の検索結果に基づいて、前記音素環境を変更した第 3 音素を生成し、該第 3 音素に対応する音素片データを前記データベースより再度検索する再検索工程のプログラムコードと、

前記検索工程あるいは前記再検索工程による検索結果と、前記第 2 音素あるいは前記第 3 音素を対応づけてテーブルに登録する登録工程のプログラムコードとを備えることを特徴とするコンピュータ可読メモリ。

【請求項 24】 データベースで管理されている音素片データを用いて音声合成を行う音声合成装置の制御のプログラムコードが格納されたコンピュータ可読メモリであって、

前記データベース中に存在する音素片データの位置を示す位置情報と、該音素

片データに対応づけられた音素環境を考慮した音素とを対応づけて管理するテーブルを記憶する記憶工程のプログラムコードと、

合成対象の音素群の各音素環境情報とその基本周波数を獲得し、獲得された基本周波数の平均を算出する算出工程のプログラムコードと、

前記音素環境情報に対応する音素群を前記テーブルより検索する検索工程のプログラムコードと、

前記算出工程で算出された基本周波数の平均に基づいて、前記検索工程で検索された音素群から所定の音素に対応する音素片データの位置情報を前記テーブルより取得する取得工程のプログラムコードと、

前記取得工程で取得された位置情報が示す音素片データを前記データベースより取得し、その取得された音素片データの韻律を変更する変更工程のプログラムコードと

を備えることを特徴とするコンピュータ可読メモリ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音素片データを管理するデータベースを有し、そのデータベースで管理されている音素片データを用いて音声合成を行う音声合成装置及びその制御方法、コンピュータ可読メモリに関するものである。

【0002】

【従来の技術】

従来より、音声合成方法として、波形編集方式による合成方法が存在する。波形編集合成法では、1～数ピッチ分の波形素片を所望のピッチ間隔に合わせて貼り合わせるピッチ同期波形重畳法によって韻律の変更を行う。波形編集合成法では、パラメータ方式による合成法に対して、より自然な合成音声を得られる反面、韻律変更に対する許容範囲が狭いという問題がある。

【0003】

そこで、様々なバリエーションの音声データを用意し、それらを適切に選択して用いることで音質向上が図られる。音声データの選択基準としては、音素環境

(合成対象となる当該音素あるいはその両側数音素)や基本周波数F0等の情報が用いられる。

【0004】

【発明が解決しようとする課題】

しかしながら、上記従来の音声合成方法は、以下のような問題点があった。

【0005】

例えば、合成対象となる音素環境を満たす音声データが存在しない場合、音素環境に関する条件を緩めて必要とする音声データを再探索することになる。そして、この再探索を音声合成時に行うと処理が複雑になり、処理時間が増大してしまうという問題点があった。また、音声データの選択規準に基本周波数F0を用いた場合、合成対象の音声データの基本周波数F0に最も合致する音声データを得るために、各音声データについて基本周波数F0の評価を行わなくてはならなかった。

【0006】

本発明は上記問題点に鑑みてなされたものであり、音声合成を精度良く高速に行うことができる音声合成装置及びその制御方法、コンピュータ可読メモリを提供することを目的とする。

【0007】

【課題を解決するための手段】

上記の目的を達成するための本発明による音声合成装置は以下の構成を備える。即ち、

音素片データを管理するデータベースを有する音声合成装置であって、
検索対象の第1音素に対し音素環境を考慮した第2音素を生成する生成手段と

、
前記第2音素に対応する音素片データを前記データベースより検索する検索手段と、

前記検索手段の検索結果に基づいて、前記音素環境を変更した第3音素を生成し、該第3音素に対応する音素片データを前記データベースより再度検索する再検索手段と、

前記検索手段あるいは前記再検索手段による検索結果と、前記第2音素あるいは前記第3音素を対応づけてテーブルに登録する登録手段とを備える。

【0008】

また、好ましくは、前記登録手段は、前記検索手段あるいは前記再検索手段によって検索された音素片データの平均基本周波数を計算する計算手段と、

前記計算手段で計算された平均基本周波数に基づいて、前記検索された音素片データ群を整列する整列手段とを備え、

前記整列手段で整列された音素片データ群の順番で、前記音素片データ群と前記第2音素あるいは第3音素を対応づけて前記テーブルに登録する。

【0009】

また、好ましくは、前記第2音素は、前記第1音素の左右音素の音素環境を考慮したトライホンである。

【0010】

また、好ましくは、前記第3音素は、前記第1音素の左右音素どちらかあるいはその両方の音素環境を考慮した音素である。

【0011】

また、好ましくは、前記第3音素は、前記第1音素が母音の場合には該第1音素の左音素環境を考慮した音素、子音の場合には該第1音素の右音素環境を考慮した音素である。

【0012】

また、好ましくは、前記登録手段は、更に、前記検索された音素片データの平均基本周波数を量子化する量子化手段を備える。

【0013】

また、好ましくは、前記計算手段は、前記量子化手段で量子化された音素片データ群の各平均基本周波数の内、対応する音素片データが存在しないものについては、その近傍の平均基本周波数で対応する音素片データが存在する平均基本周波数を用いて補間する。

【0014】

上記の目的を達成するための本発明による音声合成装置は以下の構成を備える。
即ち、

前記データベース中に存在する音素片データの位置を示す位置情報と、該音素片データに対応づけられた音素環境を考慮した音素とを対応づけて管理するテーブルを記憶する記憶手段と、

合成対象の音素群の各音素環境情報とその基本周波数を獲得し、獲得された基本周波数の平均を算出する算出手段と、

前記音素環境情報に対応する音素群を前記テーブルより検索する検索手段と、

前記算出手段で算出された基本周波数の平均に基づいて、前記検索手段で検索された音素群から所定の音素に対応する音素片データの位置情報を前記テーブルより取得する取得手段と、

前記取得手段で取得された位置情報が示す音素片データを前記データベースより取得し、その取得された音素片データの韻律を変更する変更手段と

を備える。

【0015】

また、好ましくは、前記変更手段による韻律の変更は、ピッチ同期波形重畳法を用いる。

【0016】

また、好ましくは、前記音素環境を考慮した音素の基本周波数が量子化されている場合、前記記憶手段は、その量子化された基本周波数と、該音素に対応する音素片データが存在する前記データベース中の位置を示す位置情報とを対応づけて前記テーブルに管理する。

【0017】

また、好ましくは、前記音素環境を考慮した音素の基本周波数が量子化されている場合、前記算出手段は、合成対象の音素群の各音素環境情報を獲得し、また、その量子化された音素群の各基本周波数の平均を算出する。

【0018】

上記の目的を達成するための本発明による音声合成装置の制御方法は以下の構成を備える。即ち、

音素片データを管理するデータベースを有する音声合成装置の制御方法であって、

検索対象の第1音素に対し音素環境を考慮した第2音素を生成する生成工程と、

前記第2音素に対応する音素片データを前記データベースより検索する検索工程と、

前記検索工程の検索結果に基づいて、前記音素環境を変更した第3音素を生成し、該第3音素に対応する音素片データを前記データベースより再度検索する再検索工程と、

前記検索工程あるいは前記再検索工程による検索結果と、前記第2音素あるいは前記第3音素を対応づけてテーブルに登録する登録工程とを備える。

【0019】

上記の目的を達成するための本発明による音声合成装置の制御方法は以下の構成を備える。即ち、

データベースで管理されている音素片データを用いて音声合成を行う音声合成装置の制御方法であって、

前記データベース中に存在する音素片データの位置を示す位置情報と、該音素片データに対応づけられた音素環境を考慮した音素とを対応づけて管理するテーブルを記憶する記憶工程と、

合成対象の音素群の各音素環境情報とその基本周波数を獲得し、獲得された基本周波数の平均を算出する算出工程と、

前記音素環境情報に対応する音素群を前記テーブルより検索する検索工程と、

前記算出工程で算出された基本周波数の平均に基づいて、前記検索工程で検索された音素群から所定の音素に対応する音素片データの位置情報を前記テーブルより取得する取得工程と、

前記取得工程で取得された位置情報が示す音素片データを前記データベースより取得し、その取得された音素片データの韻律を変更する変更工程とを備える。

【0020】

上記の目的を達成するための本発明によるコンピュータ可読メモリは以下の構成を備える。即ち、

音素片データを管理するデータベースを有する音声合成装置の制御のプログラムコードが格納されたコンピュータ可読メモリであって、

検索対象の第1音素に対し音素環境を考慮した第2音素を生成する生成工程のプログラムコードと、

前記第2音素に対応する音素片データを前記データベースより検索する検索工程のプログラムコードと、

前記検索工程の検索結果に基づいて、前記音素環境を変更した第3音素を生成し、該第3音素に対応する音素片データを前記データベースより再度検索する再検索工程のプログラムコードと、

前記検索工程あるいは前記再検索工程による検索結果と、前記第2音素あるいは前記第3音素を対応づけてテーブルに登録する登録工程のプログラムコードとを備える。

【0021】

上記の目的を達成するための本発明によるコンピュータ可読メモリは以下の構成を備える。即ち、

前記データベース中に存在する音素片データの位置を示す位置情報と、該音素片データに対応づけられた音素環境を考慮した音素とを対応づけて管理するテーブルを記憶する記憶工程のプログラムコードと、

合成対象の音素群の各音素環境情報とその基本周波数を獲得し、獲得された基本周波数の平均を算出する算出工程のプログラムコードと、

前記音素環境情報に対応する音素群を前記テーブルより検索する検索工程のプログラムコードと、

前記算出工程で算出された基本周波数の平均に基づいて、前記検索工程で検索された音素群から所定の音素に対応する音素片データの位置情報を前記テーブルより取得する取得工程のプログラムコードと、

前記取得工程で取得された位置情報が示す音素片データを前記データベースよ

り取得し、その取得された音素片データの韻律を変更する変更工程のプログラムコードと

を備える。

【0022】

【発明の実施の形態】

以下、図面を参照して本発明の好適な一実施形態を詳細に説明する。

【0023】

<実施形態1>

図1は本発明の実施形態1の音声合成装置の構成を示す図である。

【0024】

103はCPUであり、本発明で実行される数値演算・制御及び各種構成要素の制御等の処理を行う。102はRAMであり、本発明で実行される処理のワークエリア、各種データの一時退避領域である。101はROMであり、本発明で実行される処理のプログラム等の各種制御プログラムを格納している。また、音声合成に用いるための音素片データを管理するデータベース101aを格納する領域を有している。109は外部記憶装置であり、処理されたデータを記憶する領域として機能する。105はD/A変換器であり、当該音声合成処理装置で合成されたデジタル音声データをアナログ音声データに変換して、スピーカ110で出力する。

【0025】

106は表示制御部であり、当該音声合成処理装置の処理状態や処理結果、ユーザインタフェースをディスプレイ111に表示する際の制御を行う。107は入力制御部であり、キーボード112から入力されたキー情報を認識して指示された処理を実行する。108は通信制御部であり、通信ネットワーク113を介してデータの送受信を制御する。104はバスであり、当該音声合成装置の各種構成要素を相互に接続する。

【0026】

次に、実施形態1で実行される処理の内、処理対象の音素を検索する検索処理について、図2を用いて説明する。

【0027】

図2は本発明の実施形態1で実行される検索処理を示すフローチャートである。

【0028】

尚、実施形態1では、音素環境として各音素の両側1音素、つまり、右音素環境及び左音素環境の音素、即ち、トライホンを用いる。

【0029】

まず、ステップS1で、データベース101aから検索対象の音素pをトライホンp_{tr}に初期化する。次に、ステップS2で、データベース101aより音素pを検索する。即ち、音素pを示すラベルpが付与されている音素片データを検索する。次に、ステップS4で、データベース101a中に音素pがあるか否かを判定する。音素pがない場合（ステップS4でNO）、ステップS3に進み、音素pよりも音素環境依存度を減少させた代替音素に変更する。例えば、トライホンp_{tr}に合致する音素pがデータベース101a中に存在しなければ、右音素環境依存の音素に変更し、右音素環境依存で合致しなければ左音素環境依存の音素に変更する。また、左音素環境依存で合致しなければ音素環境とは独立に音素pを別の音素に変更するといった方法がある。あるいは、母音については左音素環境の音素を優先し、子音については右音素環境の音素を優先しても良い。また、トライホンp_{tr}に一致する音素pが存在しないとき、左あるいは右あるいはその両方の音素環境を、類似の音素環境で代用しても良い。例えば、右音素環境が'p'（パ行の子音）のとき、代替として'k'（カ行の子音）を用いても良い。このようにして、検索条件である音素pを変更した後、ステップS2に戻る。

【0030】

一方、音素pがある場合（ステップS4でYES）、ステップS5に進み、検索された音素pの各音素片データについて、平均F0（平均F0：音素片データの開始から終了までの基本周波数の平均）を計算する。尚、この計算は、対数F0（F0：時刻の関数）について行っても良いし線形F0について行っても良い。また、無声音については平均F0を0としても良いし、音素pの両側の音素の音素片データの平均F0から何らかの方法で推定しても良い。

【0031】

次に、ステップS6で、計算された平均F0を基にして、検索された各音素片データを整列（ソート）する。次に、ステップS7で、整列された音素片データをトライホンp trに対応させて登録する。登録の結果、作成される音素片データとトライホンの対応を示すインデックスは、例えば、図3のようになる。また、図3に示すように、トライホン（triphone）に対応づけて管理されるポインタ（pointer）には、その音素片データがデータベース101a中に存在する位置を示す「素片位置」とその平均F0を対応づけた表として管理される。

【0032】

以上、ステップS1～ステップS7の各ステップを、考えられるすべてのトライホンについて繰り返し、ステップS8で、全てのトライホンについて処理が終了したか否かを判定する。終了していない場合（ステップS8でNO）、ステップS1に戻る。一方、終了した場合（ステップS8でYES）、処理を終了する。

【0033】

次に、図2で説明した処理によって作成されたインデックスを用いて、合成対象の音素の音素片データを検索し音声合成を行う音声合成処理について、図4を用いて説明する。

【0034】

図4は本発明の実施形態1で実行される音声合成処理を示すフローチャートである。

【0035】

尚、音声合成処理を行うにあたり入力として、合成対象となる音素pのトライホンp tr、平均F0の軌跡が与えられる。そして、これらを基に、音素の音素片データを検索し波形重畳法により音声合成する。

【0036】

まず、ステップS9で、合成対象の音素群の平均F0の平均値F0'を求める。次に、ステップS10で、図3に示すインデックスから音素pのトライホンp trに対応する音素片データの素片位置を管理する表を検索する。例えば、トライホ

ンptrが“a. A. b”であるときには、図3より図5に示される表が得られる。尚、上記検索処理により、あらかじめ妥当な代替音素が求められているため、本ステップの結果が空になることはない。

【0037】

次に、ステップS11で、ステップS10で得られた表を基に、平均値 F_0' に最も近い平均 F_0 を持つ音素片データの素片位置を得る。ここでは、上記検索処理により、平均 F_0 に基づいて音素片データがソートされているため、探索には2分探索などの手法を用いることが可能である。次に、ステップS12で、ステップS11で得られた素片位置から音素片データをデータベース101aから取り出す。次に、ステップS13で、波形重畳法を用いてステップS12で得られた音素片データの韻律を変更する。

【0038】

以上説明したように、実施形態1によれば、考えられる全ての音素環境に対して予め音素片データの有無を確認し、音素片データが存在しない場合にはあらかじめ代替音素を用意しておくことにより、処理が単純化され高速化が図られる。また、各音素環境について存在する音素片データの平均 F_0 に関する情報をあらかじめ抽出して、それに基づいて音素片データを管理しておくので音声合成時の処理の高速化が図られる。

【実施形態2】

上記実施形態1において、図2に示したステップS5の代わりにステップS14を設け、連続的な音素片データの平均 F_0 を計算する代わりに、音素片データの平均 F_0 を量子化しても良い。この場合の処理について、図6を用いて説明する。

【0039】

図6は本発明の実施形態2で実行される検索処理を示すフローチャートである。

【0040】

尚、実施形態1の図2と同じ処理については、同じステップ番号を付加し、その詳細は省略する。

【0041】

ステップS14で、検索された音素pの各音素片データの平均F0を量子化して、量子化平均F0を得る（量子化平均F0：連続量である平均F0を適当な間隔で量子化したもの）。尚、この計算は、対数F0について行っても良いし線形F0について行っても良い。また、無声音については平均F0を0としても良いし、両側の音素片データの平均F0から何らかの方法で推定しても良い。

【0042】

次に、ステップS6aで、計算された平均F0を基にして、検索された各音素片データを整列（ソート）する。次に、ステップS7aで、整列された音素片データをトライホンptrに対応させて登録する。登録の結果、作成される音素片データとトライホンの対応を示すインデックスは、例えば、図7のようになる。また、図7に示すように、トライホン（triphone）に対応づけて管理されるポインタ（pointer）には、その音素片データがデータベース101a中に存在する位置を示す「素片位置」とその平均F0を対応づけた表として管理される。

【0043】

以上、ステップS1～ステップS7aの各ステップを、考えられるすべてのトライホンについて繰り返し、ステップS8aで、全てのトライホンについて処理が終了したか否かを判定する。終了していない場合（ステップS8aでNO）、ステップS1に戻る。一方、終了した場合（ステップS8aでYES）、処理を終了する。

【0044】

以上説明したように、実施形態2によれば、実施形態1で説明した効果に加えて、音素片データの量子化平均F0を用いることにより、音素片数の削減、検索時の計算量を減少させる効果を得ることが可能である。

〔実施形態3〕

上記実施形態2において、整列された音素片データ間を補間した後に、各音素片データをトライホンptrに対応させて登録するようにしても良い。即ち、全ての量子化された音素片データの平均F0に対してインデックスの表中に対応する素片位置が見つかるような構成にしても良い。この場合の処理について、図8を

用いて説明する。

【0045】

図8は本発明の実施形態3で実行される検索処理を示すフローチャートである。

【0046】

尚、実施形態2の図6と同じ処理については、同じステップ番号を付加し、その詳細は省略する。

【0047】

ステップS15で、整列された音素片データ間を補間する。ステップS7bで、補間された音素片データをトライホンptrに対応させて登録する。登録の結果、作成される音素片データとトライホンの対応を示すインデックスは、例えば、図9のようになる。また、図9に示すように、トライホン(triphone)に対応づけて管理されるポインタ(pointer)には、その音素片データがデータベース101a中に存在する位置を示す「素片位置」とその平均F0を対応づけた表として管理される。

【0048】

以上、ステップS1～ステップS7bの各ステップを、考えられるすべてのトライホンについて繰り返し、ステップS8bで、全てのトライホンについて処理が終了したか否かを判定する。終了していない場合(ステップS8bでNO)、ステップS1に戻る。一方、終了した場合(ステップS8bでYES)、処理を終了する。

【0049】

以上説明したように、実施形態3によれば、実施形態2で得られる効果に加えて、すべての音素片データの素片位置を管理しているので、図4のステップS11で説明した処理を、単なる表参照として実現することができ、処理を簡略化することができる。

【0050】

尚、本発明は、複数の機器(例えばホストコンピュータ、インタフェイス機器、リーダ、プリンタなど)から構成されるシステムに適用しても、一つの機器が

らなる装置（例えば、複写機、ファクシミリ装置など）に適用してもよい。

【0051】

また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

【0052】

この場合、記憶媒体から読出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

【0053】

プログラムコードを供給するための記憶媒体としては、例えば、フロッピディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモ리카ード、ROMなどを用いることができる。

【0054】

また、コンピュータが読出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOS（オペレーティングシステム）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0055】

更に、記憶媒体から読出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0056】

【発明の効果】

【0057】

以上説明したように、本発明によれば、音声合成を精度良く高速に行うことができる音声合成装置及びその制御方法、コンピュータ可読メモリを提供できる。

【図面の簡単な説明】

【図1】

本発明の実施形態1の音声合成装置の構成を示す図である。

【図2】

本発明の実施形態1で実行される検索処理を示すフローチャートである。

【図3】

本発明の実施形態1で管理されるインデックスを示す図である。

【図4】

本発明の実施形態1で実行される音声合成処理を示すフローチャートである。

【図5】

本発明の実施形態1で管理されるインデックスより得られる表を示す図である。

【図6】

本発明の実施形態2で実行される検索処理を示すフローチャートである。

【図7】

本発明の実施形態2で管理されるインデックスを示す図である。

【図8】

本発明の実施形態3で実行される検索処理を示すフローチャートである。

【図9】

本発明の実施形態3で管理されるインデックスを示す図である。

【符号の説明】

101 ROM

101a データベース

102 RAM

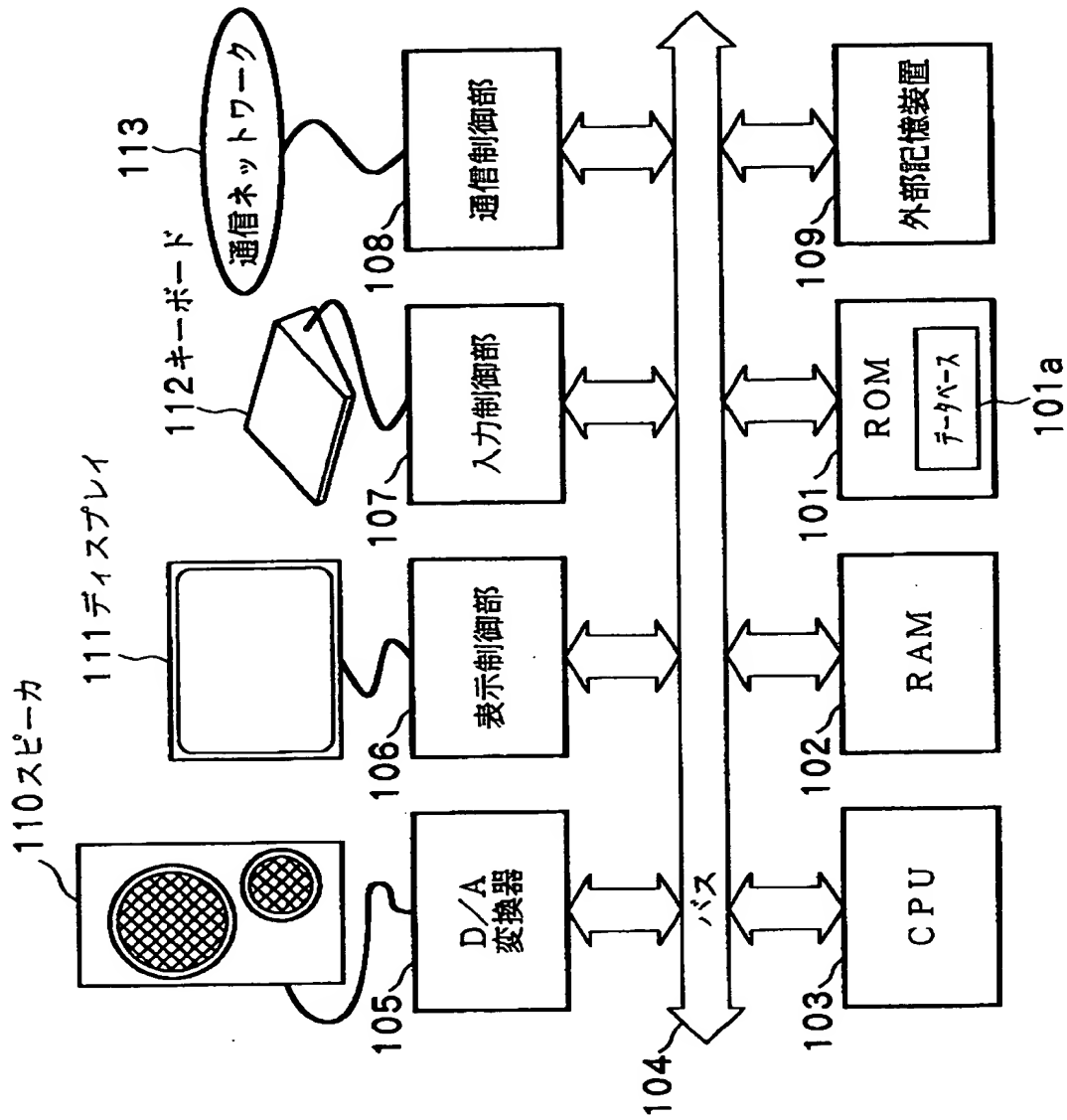
103 CPU

104 バス

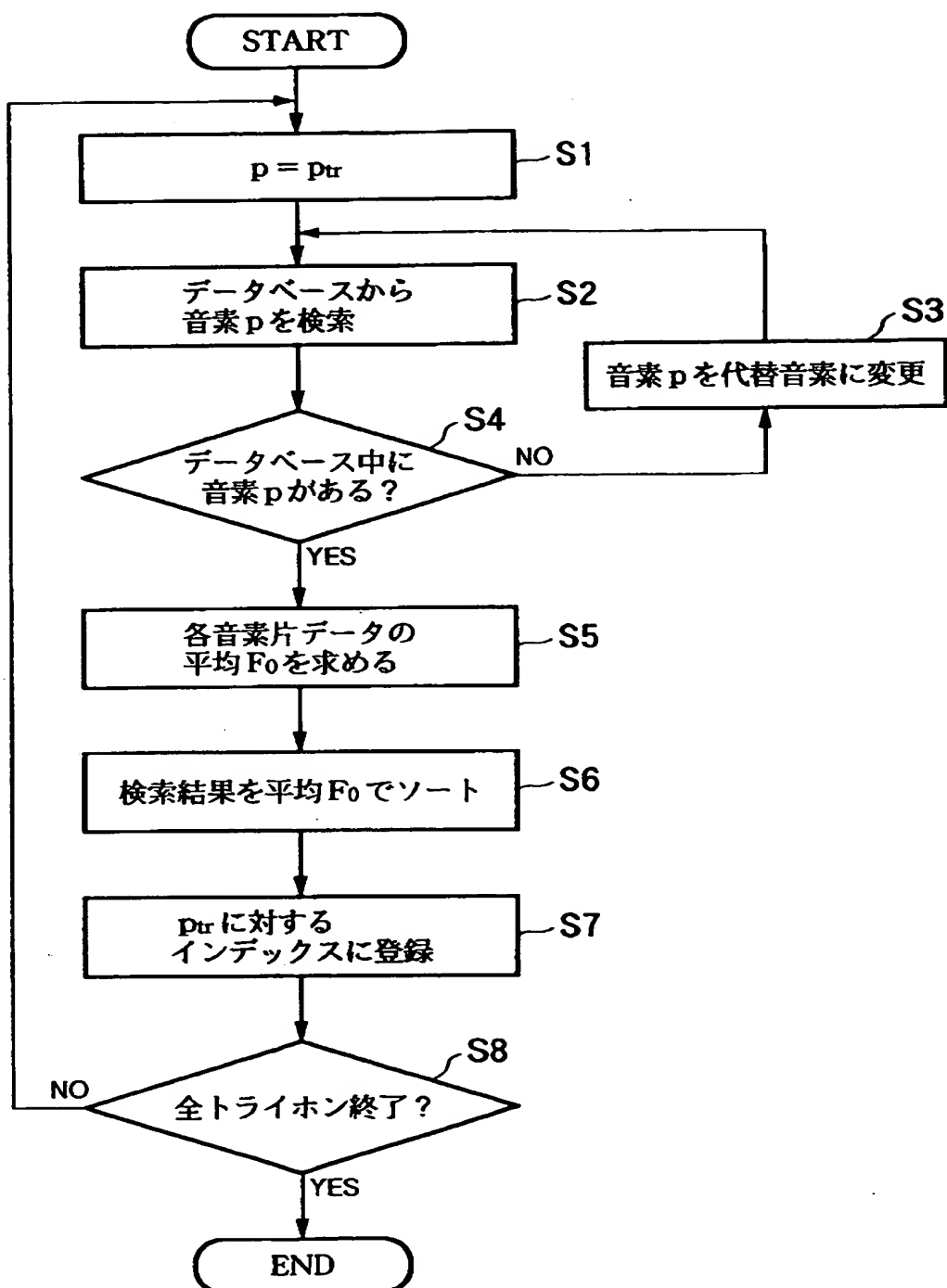
- 105 D/A変換器
- 106 表示制御部
- 107 入力制御部
- 108 通信制御部
- 109 外部記憶装置
- 110 スピーカ
- 111 ディスプレイ
- 112 キーボード
- 113 通信ネットワーク

【書類名】 図面

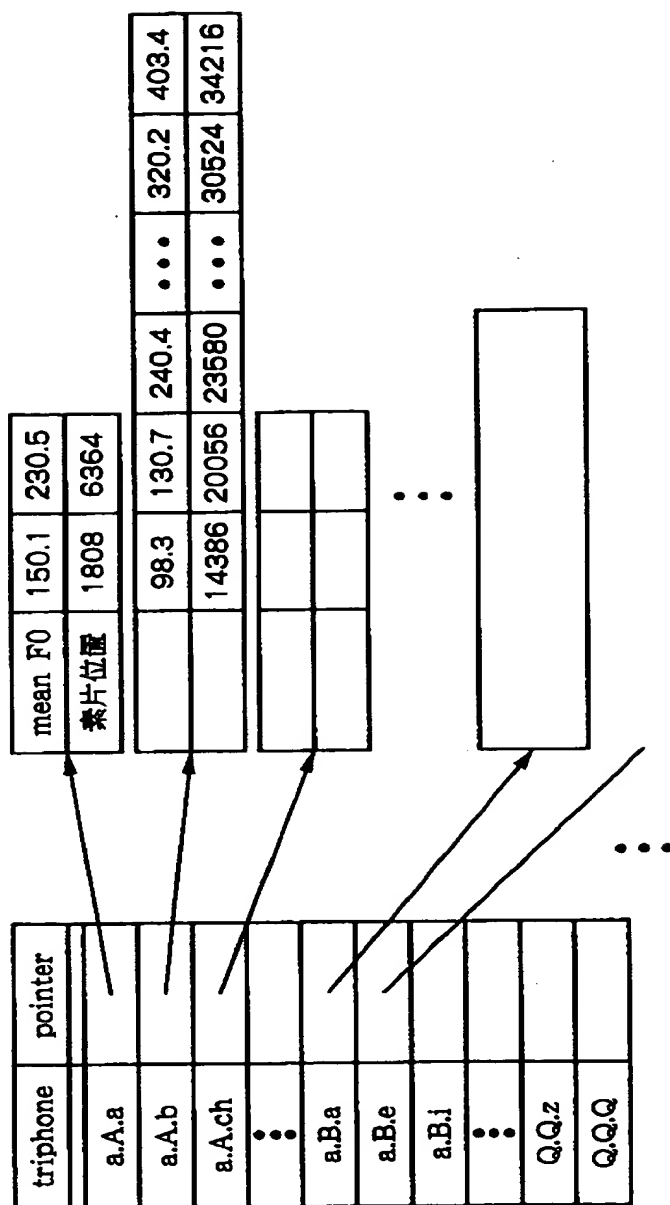
【図 1】



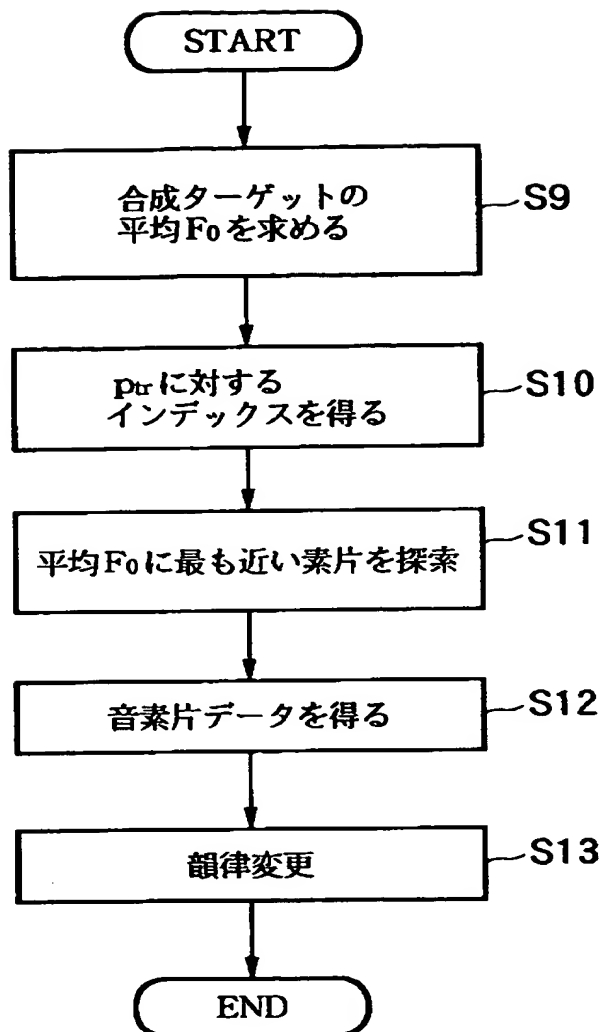
【図 2】



【图 3】



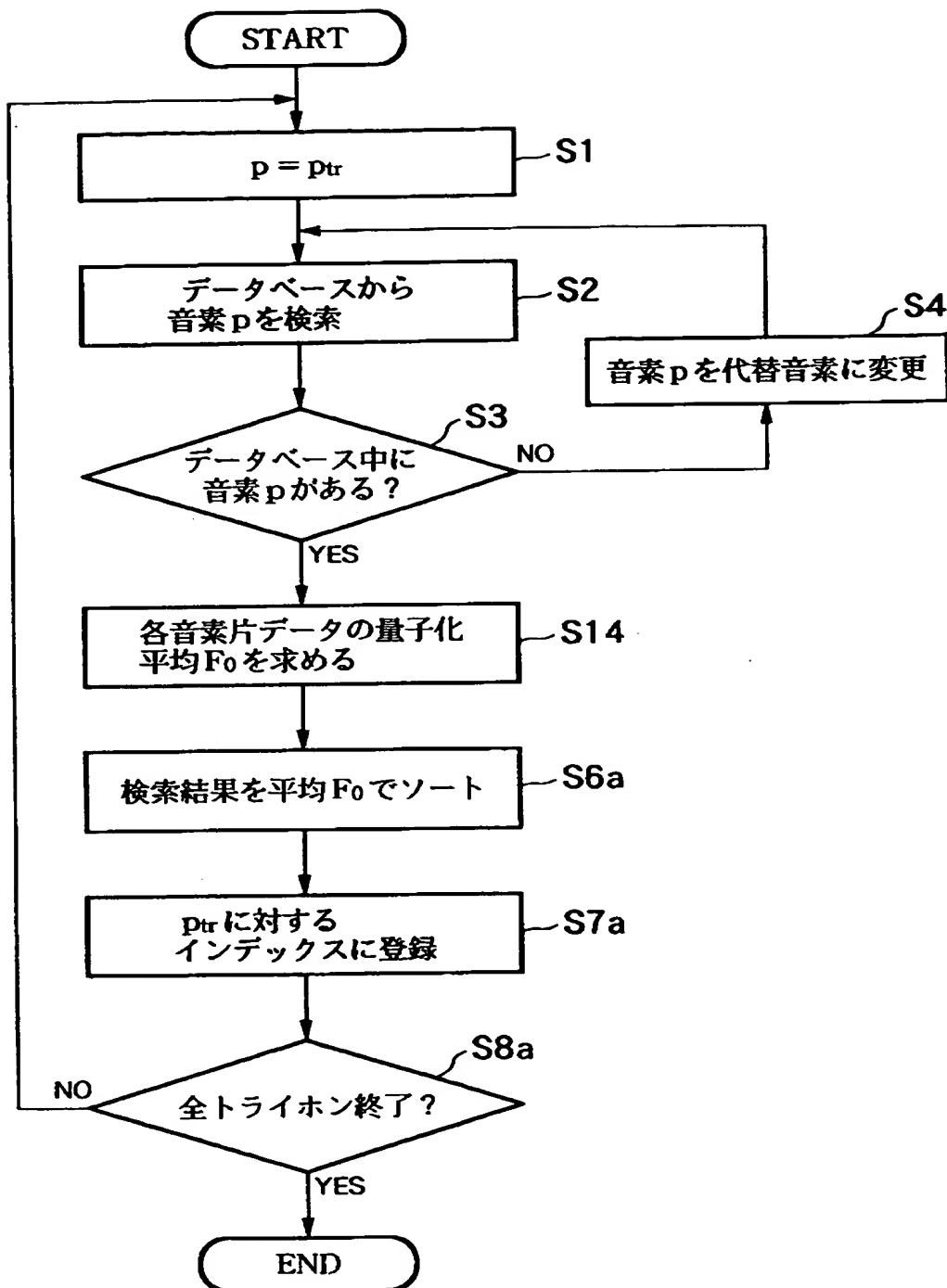
【図 4】



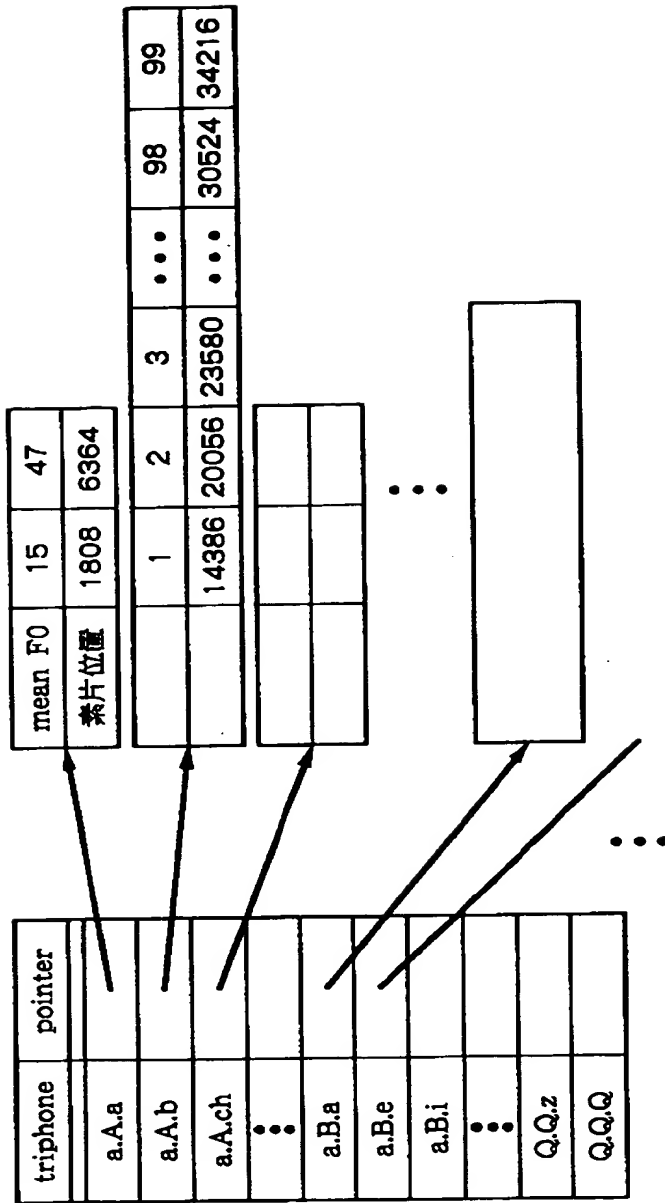
【図 5】

mean F0	98.3	130.7	240.4	...	320.2	403.4
素片位置	14386	20056	23580	...	30524	34216

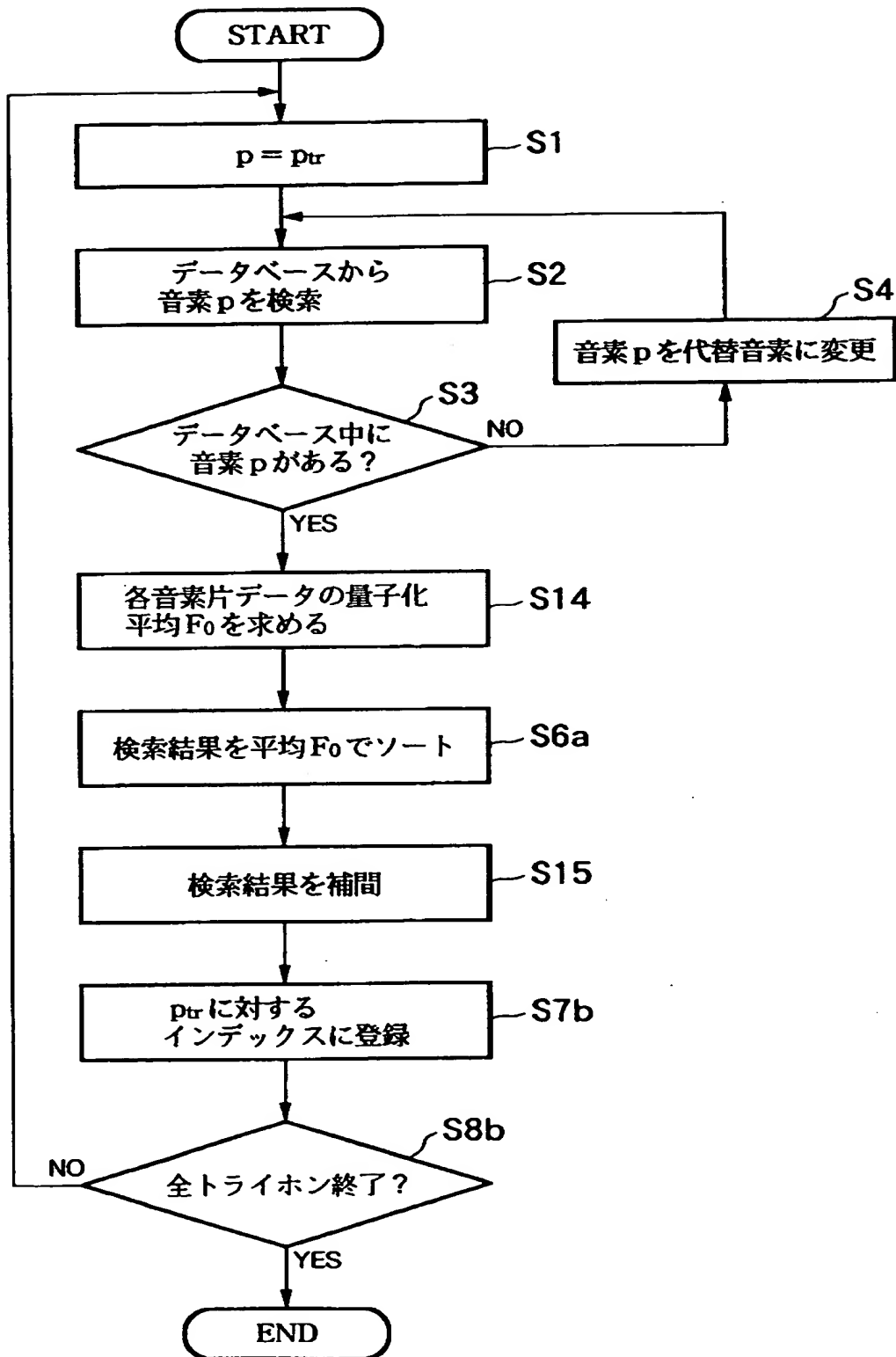
【図 6】



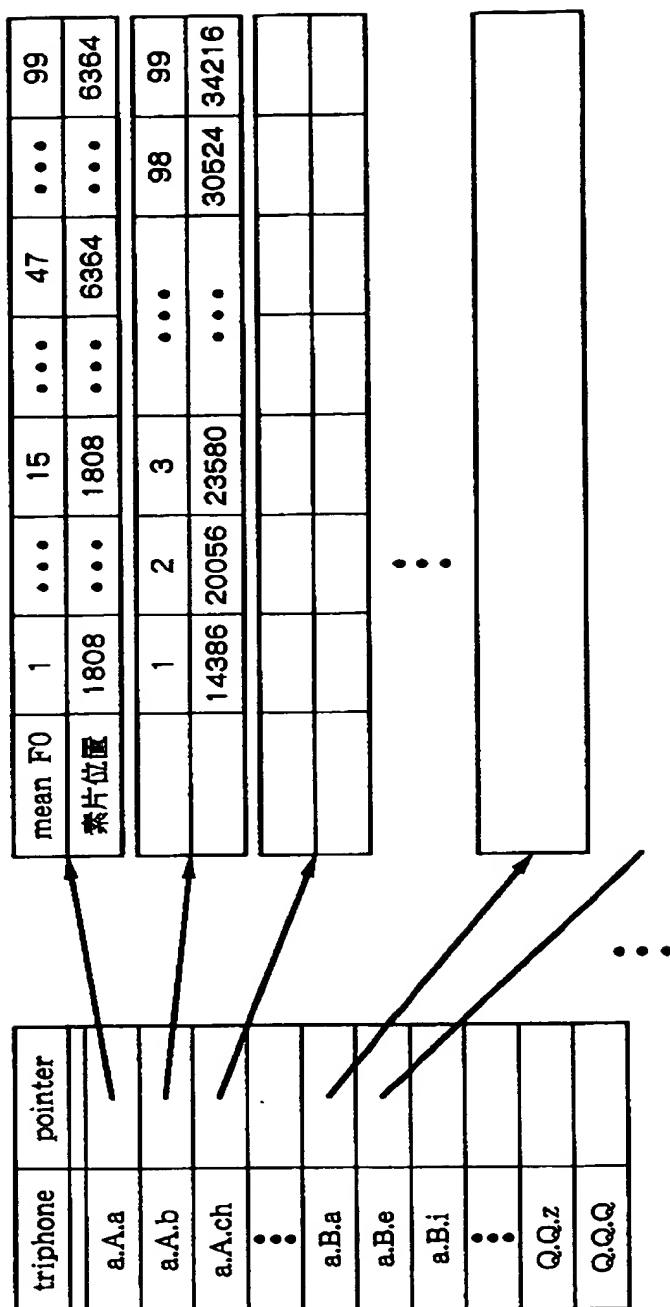
【図 7】



【図 8】



【图 9】



【書類名】 要約書

【課題】 音声合成を精度良く高速に行うことができる音声合成装置及びその制御方法、コンピュータ可読メモリを提供する。

【解決手段】 検索対象の第1音素に対し音素環境を考慮した第2音素を生成する。その第2音素に対応する音素片データをデータベース101aより検索する。その検索結果に基づいて、音素環境を変更した第3音素を生成し、該第3音素に対応する音素片データをデータベース101aより再度検索する。上記検索あるいは再検索による検索結果と、第2音素あるいは第3音素を対応づけてテーブルに登録する。

【選択図】 図1

【書類名】 職権訂正データ
【訂正書類】 特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】 000001007
【住所又は居所】 東京都大田区下丸子3丁目30番2号
【氏名又は名称】 キヤノン株式会社

【代理人】 申請人

【識別番号】 100076428
【住所又は居所】 東京都千代田区麴町5丁目7番地 紀尾井町TBR
ビル507号室
【氏名又は名称】 大塚 康德

【選任した代理人】

【識別番号】 100093908
【住所又は居所】 東京都千代田区麴町5丁目7番地 紀尾井町TBR
ビル507号室
【氏名又は名称】 松本 研一

【選任した代理人】

【識別番号】 100101306
【住所又は居所】 東京都千代田区麴町5丁目7番地 紀尾井町TBR
ビル507号室
【氏名又は名称】 丸山 幸雄

出 願 人 履 歴 情 報

識別番号 [000001007]

1. 変更年月日 1990年 8月30日
[変更理由] 新規登録
住 所 東京都大田区下丸子3丁目30番2号
氏 名 キヤノン株式会社